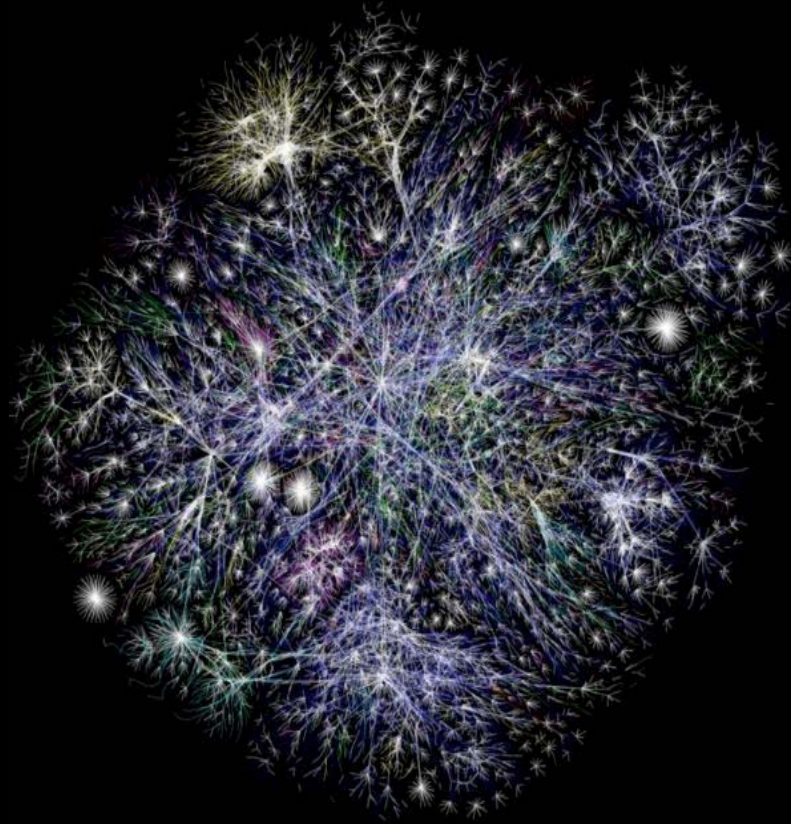


The Advent of the Internet



SIMS 103

Nov. 19, 2005

Geoff Nunberg

A Convergence of Causes

Technological developments

Policy choices

Economic forces

Social responses

Technological Beginnings

1969: ARPA (Advanced Research Projects Agency of DOD) (later DARPA) creates Arpanet, linking time-sharing computers at four research sites by telephone lines. Net makes use of packet-switching, rather than circuit switching, as with phone communication at the time.

1971: File Transfer Protocol (FTP) permits easy exchange of files between sites.

1974 Bob Kahn and Vin Cerf ("Father of the Internet") demonstrate Transfer Control Protocol, which enables machines to route & assemble data packets.

1974: Ethernet developed at Xerox Palo Alto Research Center (PARC), allowing communication among machines on local networks.



Arpanet 1971

Internet Development

1980's: NSF funds national backbone to connect computer research centers.

Other gov't-funded networks (BITNET, CSNET) emerge

1980's: Commercial networks begin to emerge

1983: Domain Name System (DNS) introduced to keep up with growing number of hosts, introduces domain names .com, .gov, .mil, .edu, etc.

Late 1980's: First Internet Service Providers emerge

1989: Australia, UK, Germany, Italy, etc. join Internet

1990: ARPANET shuts down

1991: NSF removes all restrictions on commercial use of Internet

1992: Internet Society (ISOC) formed, assumes responsibility for fixing standards through the Internet Engineering Task Force (IETF), by now a voluntary organization

1995: NSF discontinues support of infrastructure

1998: Internet Corporation for Assigned Names and Numbers (ICANN) established to oversee assignment of domain names and IP addresses, formerly under control of US government. But Bush administration indicates intent to retain control in 2005.

The Origins of Email

1971: First network email program created by Ray Tomlinson at Bolt, Beranek & Newman (BBN), with "USER@hostname.domain" addressing system. First message sent between 2 machines that are side-by-side.

But most email is restricted to proprietary networks until 1988, when MCI mail was linked to the Internet



Advent of the Home Computer

1975: Altair 8800 (sold as kit)
takes off, first "useful" home
computer.

Bill Gates and Paul Allen develop
BASIC program for the
machine; Gates drops out of
Harvard to create new company
called Micro-Soft.



Advent of the Home Computer

1973: Xerox Alto is first minicomputer to use graphical interface, desktop metaphor, & mouse (developed by Douglas Engelbart at SRI).

1977: Apple Computer introduces Apple II, first successful off-the-shelf home computer, with 1 MHz chip & 4k of RAM, at \$1298 retail. Becomes standard in educational use & business; VisiCalc is first computer spreadsheet program.

1981: IBM introduces PC, using OS developed by Microsoft. By 1982, compatible PC clones are on market.

1981: Osborne, first "portable" home computer



Osborne



IBM PC



Apple II



Xerox Alto

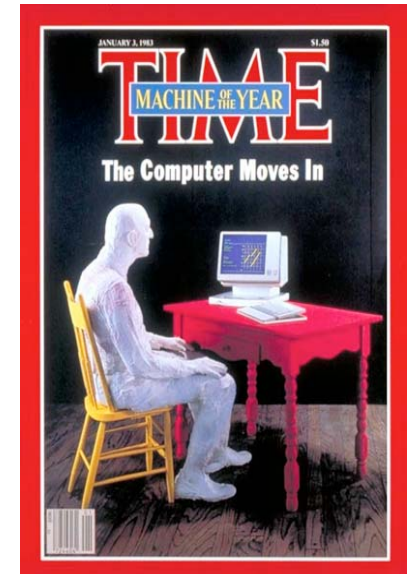
Triumph of the Home Computer

1983: Time magazine names the personal computer the "machine of the year"

1984: With famous Superbowl commercial, Apple launches Macintosh, first successful home computer to use Graphical User Interface (GUI) based on Xerox research

1985: Microsoft introduces Windows as add-on environment to MS-DOS operating system; later releases improve on system.

1980-1990: Home computer penetration increases rapidly, driven by spreadsheets, word-processing, desktop publishing, and (later) email.



Early Windows desktop



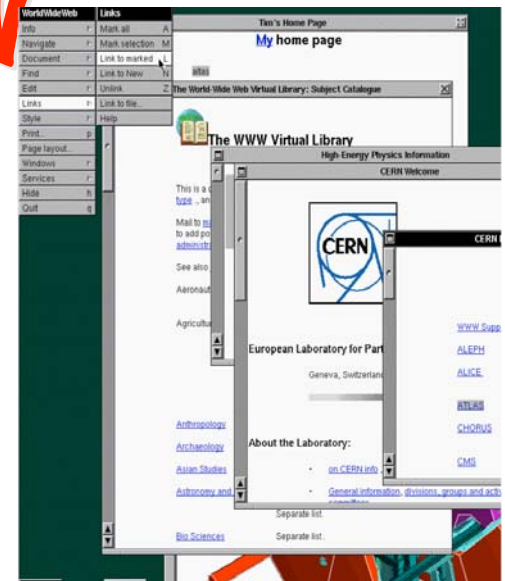
The Emergence of the WWW

1945: Vannevar Bush writes "As We May Think" in *The Atlantic*; envisions Memex machine to follow links between documents on microfiche

1965: Ted Nelson coins the term "hypertext" to describe "compound documents" formed by links among documents

1990: Tim Berners-Lee of CERN coins the term "World Wide Web"; develops HTTP protocol for transmitting hypertext documents between clients and servers and first Web browser making use of hypertext links.

Ca. 1990-: Pay-based online services like AOL, CompuServe, and Prodigy market connectivity + proprietary content (games, chat rooms, e-commerce, instant messaging etc.) to users unfamiliar with computers, first for hourly and then for monthly fee. By 1998, AOL has 15m. members.



The Growth of the WWW

1993: Mark Andreessen's Mosaic browser released by NCSA, which runs on Windows and permits easy integration of graphics in Web pages.

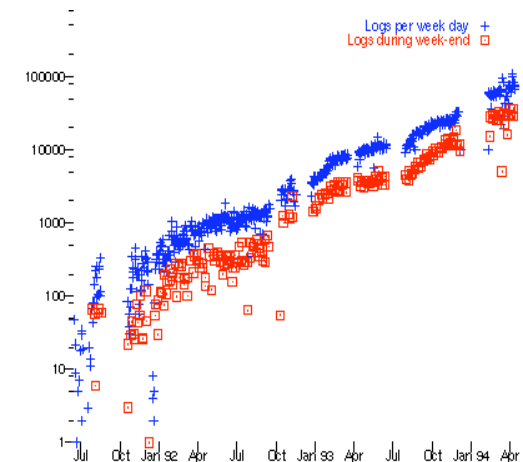
CERN announces that W3 technology will be available free to everyone.

1994: Over 200 HTTP servers; traffic on CERN server has grown 1000-fold since first launched. From the mid-90s, Internet use roughly doubles every year.

1994: Andreessen, now in private sector, releases Netscape Navigator browser.

1995: Microsoft releases Internet Explorer bundled with Windows 95 to compete with Netscape, which rapidly loses market share. Win95 also bundles dial-up software to make connection easy.

1995 AOL makes Internet available to all subscribers



The Addition of Search

- 1991: Gopher, developed at U. Minnesota, creates searchable index of FTP sites
- 1994: Infoseek and Lycos search engines launched.
- Jerry Yang and David Filo introduce Yahoo!, a directory of Web sites.
- 1995: AltaVista launched by DEC; company regards it as showpiece for its hardware
- 1997 Larry Page and Sergey Brin launch Google, which makes use of Page Rank algorithm to rank pages according to popularity.
- 1998: Goto.com (later Overture, later Yahoo! Search) introduces pay-per-click advertising

LYCOS

altavista™

infoseek™

YAHOO!

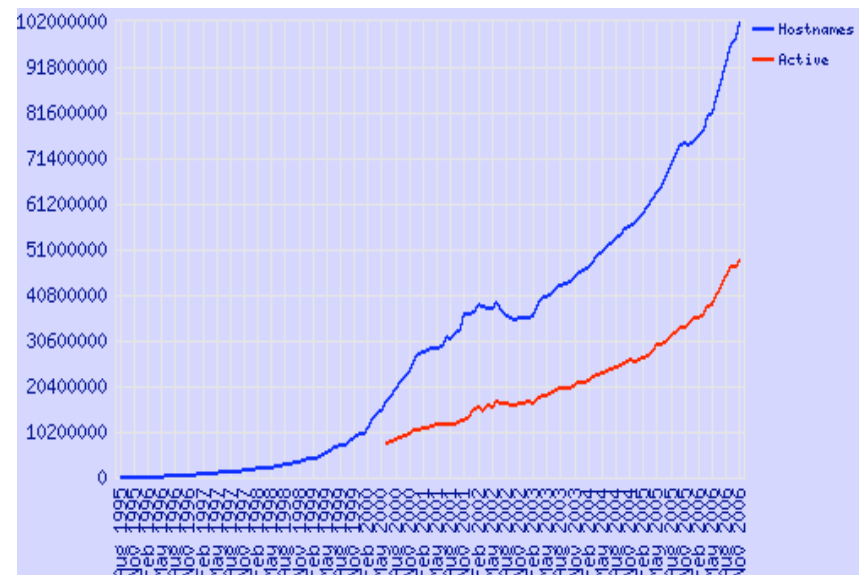


The Web Takes Off

1994-2005: Internet use increases rapidly, driven by email, E-commerce, news & information, pornography & gambling. By 2005 there are an estimated 100m Web sites.

2000- Growth of broadband enables exchange of audio & video content; blogs and social networking sites proliferate, etc.

2005: 68 percent of American adults and 90 percent of American teenagers have used the Internet.



The Internationalization of the Web

Initial prevalence of English on the Web

1997 -- Nunberg & Schuetze survey of 2.3 million pages from Web crawl: 85 percent of content is in English, including about 35 percent of content in non-English speaking domains.

Eng. Most common in N. Europe, less-developed nations

2000: Jack Xu Excite@Home; survey of around 600 million pages: 72% are in English

Further Studies

2002 study at Online Computer Library Center of 2200 random IP's

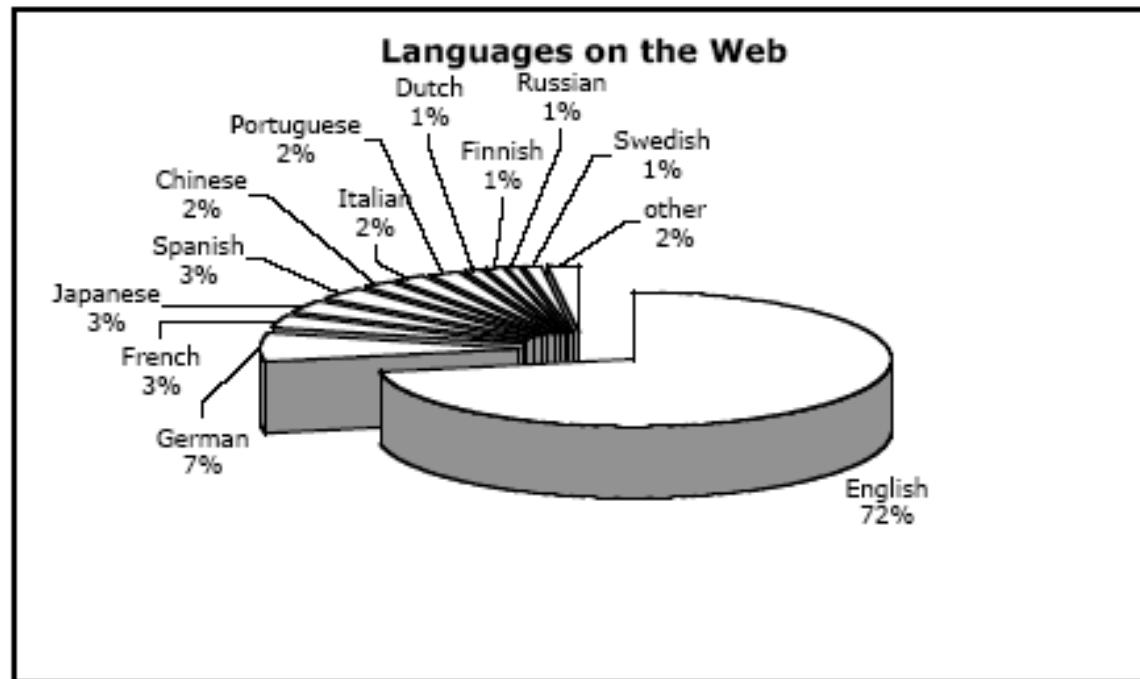


Figure 17. Proportion of languages on the web in a random sample of web pages (source: and O'Niell, Lavoie and Bennett 2003)

Fears of English Hegemony...

“The Web is the ultimate act of intellectual colonialism.” Director of Russian ISP, 1999

“Nearly 70 per cent of the world's Web sites are in English, at times crowding out voices and views.” Kofi Annan, 12-Jan-04

And Anglophone Triumphalism...

The Internet is "a great force for the Anglification of the planet."

“[Thanks to the Internet,] English will be the native language of a majority of the world by some time in the next century.” Editor, *The Futurist*

The Polyglot Web

Proportion of English is dropping sharply, aided by widespread adoption of unicode, first developed in 1991 to enable rendering of non-Roman writing systems

Relatively slower growth in N.America, Europe as opposed to Latin America, East Asia

Relative proportion of English is dropping in non-English-speaking domains as usage widens

Growth of Other Languages on the Web

Top Ten Languages Used in the Web (Number of Internet Users by Language)					
TOP TEN LANGUAGES IN THE INTERNET	% of all Internet Users	Internet Users by Language	Internet Penetration by Language	Internet Growth for Language (2000 - 2006)	World Population 2006 Estimate for the Language
English	29.7 %	322,600,837	28.7 %	135.2 %	1,125,664,397
Chinese	13.3 %	144,301,513	10.8 %	346.7 %	1,340,767,863
Japanese	7.9 %	86,300,000	67.2 %	83.3 %	128,389,000
Spanish	7.5 %	81,729,671	18.7 %	231.1 %	437,502,257
German	5.4 %	58,854,682	61.3 %	113.2 %	95,982,043
French	4.6 %	49,660,498	13.0 %	307.1 %	381,193,149
Portuguese	3.1 %	34,064,760	14.8 %	349.6 %	230,846,275
Korean	3.1 %	32,372,000	45.8 %	78.0 %	73,945,860
Italian	2.7 %	28,870,000	48.8 %	118.7 %	59,115,261
Russian	2.2 %	23,700,000	16.5 %	664.5 %	143,682,757
TOP TEN LANGUAGES	79.5 %	863,981,961	21.5 %	166.7 %	4,017,088,863
Rest of World Languages	20.5 %	222,268,942	9.0 %	500.0 %	2,482,608,197
WORLD TOTAL	100.0 %	1,086,250,903	16.7 %	200.9 %	6,499,697,060

The Polyglot Web

Clear that proportion of English is dropping

Relatively slower growth in N.America, Europe as
opposed to Latin America, East Asia

Relative proportion of English is dropping in non-
English-speaking domains

... But the *perception* of English dominance persists

One reason: English still by far the most widely
used single language...

The “Omnigooglization” of the Web

1st 50 Google hits for “Roland Barthes”:
44 English, 4 French, 1 Spanish, 1 German

The screenshot shows a Google search interface with the query "roland barthes" entered. The search results are displayed under the "Web" tab, showing the first 100 results out of approximately 341,000. The results list includes links to various websites, such as "Roland Barthes: Mythologies", "Roland Barthes", "PopCultures.com | Theorists and Critics | Roland Barthes", "Semiotics", "en.wikipedia.org/wiki/Roland_Barthes", "www2.ucsc.edu/people/tuttle/", and "ROLAND BARTHES THE DISCOURSE OF HISTORY, translated by Stephen ...". Each result includes a brief description and a link to the full page. The search bar at the top shows the query "roland barthes" and the search button. The Google logo is visible on the left. The page is titled "Web" and "Results 1 - 100 of about 341,000 for 'roland barthes'".

Web Images Groups News Froogle Local^{New!} more »

Google "roland barthes" Search Advanced Search Preferences

Web Results 1 - 100 of about 341,000 for "roland barthes" [defir]

Roland Barthes: Mythologies
Set of lectures, by Tony McNeill, on Barthes' view of postmodern literature, art, mythology, and semiology.
orac.sund.ac.uk/~os0tmc/myth.htm - 65k - [Cached](#) - [Similar pages](#)

Roland Barthes
... **Roland Barthes** was born in Cherbourg, Manche. After his father's death in a ... **Roland Barthes**: Un regard politique sur le signe by LS Calvet (1973); ...
www.kirjasto.sci.fi/rbarthes.htm - 17k - [Cached](#) - [Similar pages](#)

PopCultures.com | Theorists and Critics | Roland Barthes
Texts by and about **Roland Barthes**. ... **Roland Barthes** (abstracted from the Fifty Key Contemporary Thinkers by John Lechte); **Roland Barthes** ...
www.popcultures.com/theorists/barthes.html - 7k - [Cached](#) - [Similar pages](#)

Semiotics
... Basics of Semiotics, **Roland Barthes**, Paul Bouissac ... **Roland Barthes**; Ron Burnett; Sonesson Pictorial Semiotics; Sonesson Photography ...
www.cudenver.edu/~mryder/itc_data/semiotics.html - 91k - 9 Apr 2005 - [Cached](#) - [Similar pages](#)

en.wikipedia.org/wiki/Roland_Barthes
[Similar pages](#)

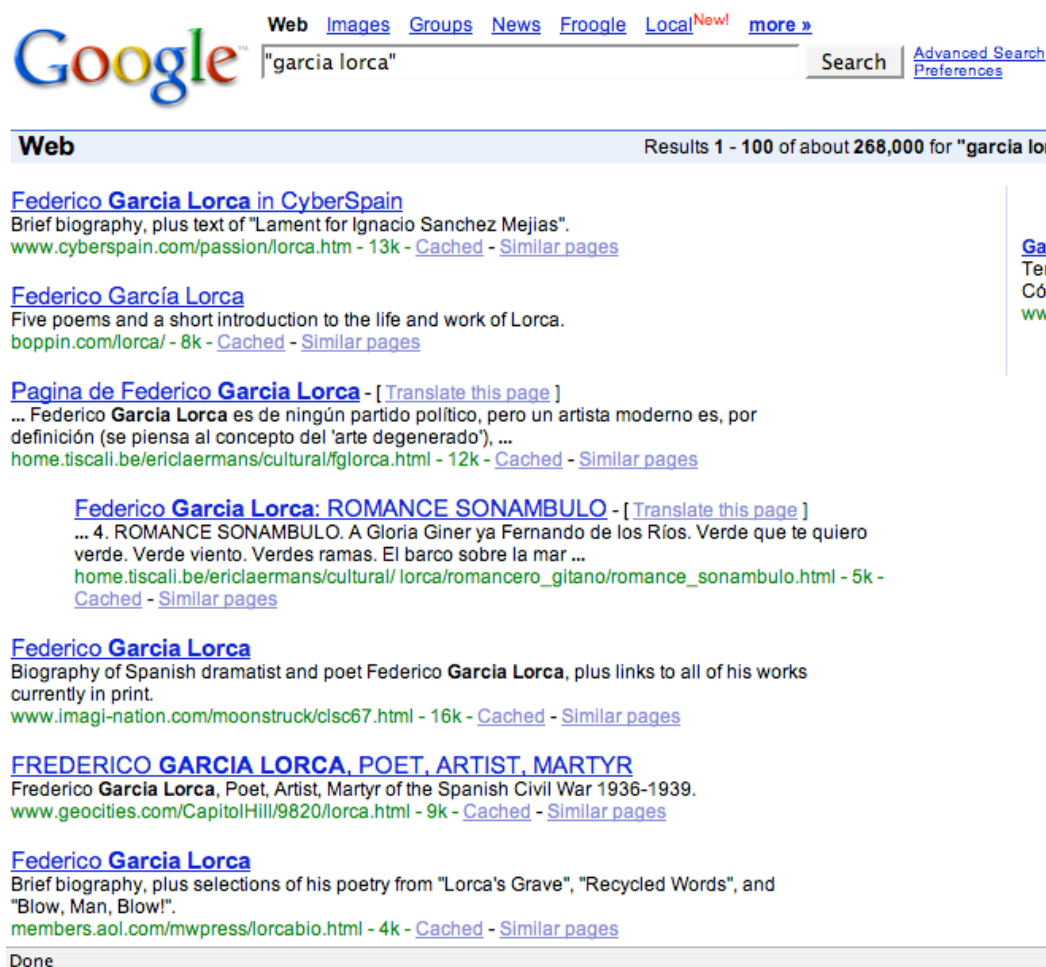
www2.ucsc.edu/people/tuttle/
[Similar pages](#)

ROLAND BARTHES THE DISCOURSE OF HISTORY, translated by Stephen ...
ROLAND BARTHES THE DISCOURSE OF HISTORY, translated by Stephen Bann. Comparative Criticism, 3 (1981): 7-20. Pagination, superscripts, and accents are not ...
www.clas.ufl.edu/users/pcraddoc/barthes.htm - 38k - [Cached](#) - [Similar pages](#)

Done

Omnigooglization, 2

1st 50 Google hits for “Garcia Lorca”:
45 English, 4 Spanish, 1 Italian



The image is a screenshot of a Google search results page for the query "garcia lorca". At the top, the Google logo is on the left, and navigation links for "Web", "Images", "Groups", "News", "Froogle", "Local", "New!", and "more »" are on the right. Below the logo is the search bar containing "garcia lorca" and a "Search" button. To the right of the search bar are links for "Advanced Search" and "Preferences". Below the search bar, a horizontal bar indicates "Web" results and "Results 1 - 100 of about 268,000 for 'garcia lo'". The search results are listed below, each with a title, a brief description, and a URL with additional information like "13k", "8k", "12k", "5k", "16k", "9k", and "4k". On the right side of the results, there is a vertical sidebar with the text "Ga", "Te", "Có", and "ww".

Google Web Images Groups News Froogle Local New! more »
"garcia lorca" Search Advanced Search Preferences

Web Results 1 - 100 of about 268,000 for "garcia lo"

Federico Garcia Lorca in CyberSpain
Brief biography, plus text of "Lament for Ignacio Sanchez Mejias".
www.cyberspain.com/passion/lorca.htm - 13k - [Cached](#) - [Similar pages](#)

Federico García Lorca
Five poems and a short introduction to the life and work of Lorca.
boppin.com/lorca/ - 8k - [Cached](#) - [Similar pages](#)

Pagina de Federico Garcia Lorca - [[Translate this page](#)]
... Federico Garcia Lorca es de ningún partido político, pero un artista moderno es, por definición (se piensa al concepto del 'arte degenerado'), ...
home.tiscali.be/ericaermans/cultural/fglorca.html - 12k - [Cached](#) - [Similar pages](#)

Federico Garcia Lorca: ROMANCE SONAMBULO - [[Translate this page](#)]
... 4. ROMANCE SONAMBULO. A Gloria Giner ya Fernando de los Ríos. Verde que te quiero verde. Verde viento. Verdes ramas. El barco sobre la mar ...
home.tiscali.be/ericaermans/cultural/lorca/romancero_gitano/romance_sonambulo.html - 5k - [Cached](#) - [Similar pages](#)

Federico Garcia Lorca
Biography of Spanish dramatist and poet Federico Garcia Lorca, plus links to all of his works currently in print.
www.imagi-nation.com/moonstruck/clsc67.html - 16k - [Cached](#) - [Similar pages](#)

FREDERICO GARCIA LORCA, POET, ARTIST, MARTYR
Federico Garcia Lorca, Poet, Artist, Martyr of the Spanish Civil War 1936-1939.
www.geocities.com/CapitolHill/9820/lorca.html - 9k - [Cached](#) - [Similar pages](#)

Federico Garcia Lorca
Brief biography, plus selections of his poetry from "Lorca's Grave", "Recycled Words", and "Blow, Man, Blow".
members.aol.com/mwpress/lorcabo.html - 4k - [Cached](#) - [Similar pages](#)

Done

The bright side of omnigooglization, 1

Ease of diffusion increases accessibility of news & information for smaller languages. Eliminates disparities in access that are inherent in older media.

The bright side of omnigooglization, 2

International distribution has major effects for languages with large international diasporas (e.g., Chinese, Russian) and for languages with large postcolonial populations (e.g., French).

Does the Internet created “global language communities” that eliminate the importance of national communities? Probably not.

Millenarian Visions

Predictions that Internet/Web will:

"Disintermediate" commerce & discourse (eliminate the middleman, like retailers & wholesalers)

e.g., retailers like Amazon, travel sites, EBay, classified ads, etc.

Cf resistance from intermediaries (like real-estate brokers), difficulties in achieving disintermediation with education, business sourcing, disintermediation of news by bloggers...

Millenarian Visions, 2

Predictions that Internet/Web will Dematerialize
content
cf online newspapers, i-Tunes.

The End of the Book?



2000 Microsoft Reader debuts

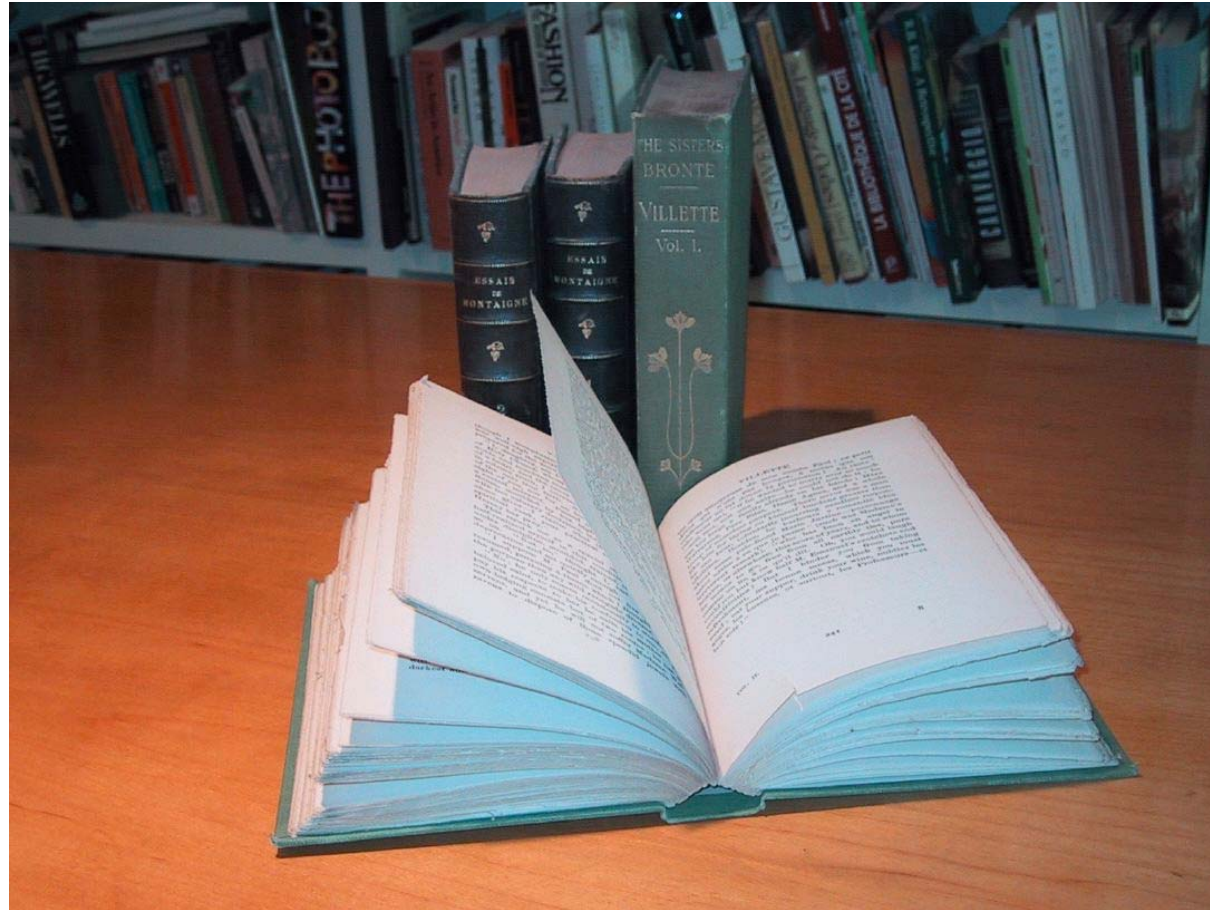
2009 E-books begin to outsell paper

2016 Last print
newspaper

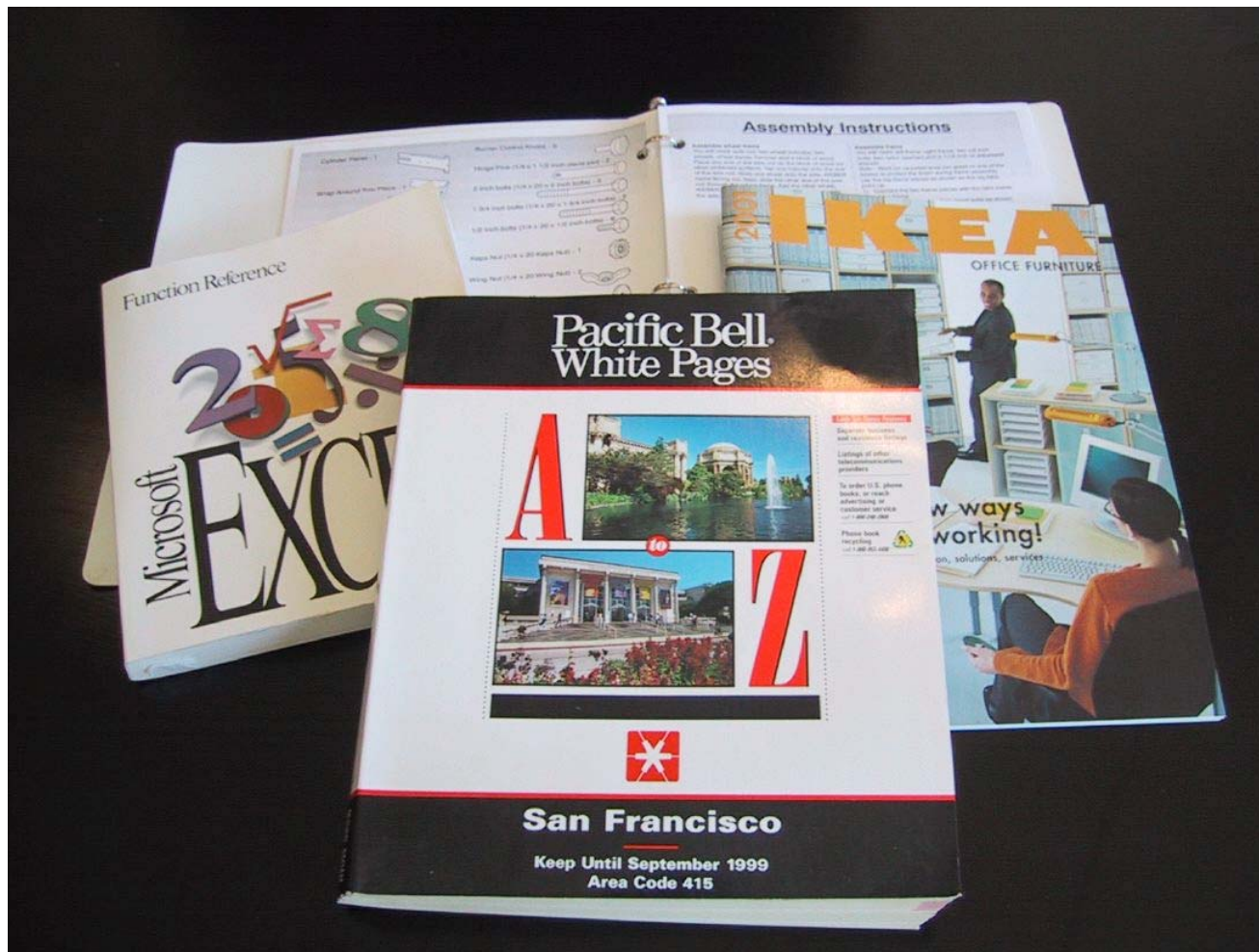
2019 Paper books remain popular among collectors. . .

2020 Dictionary
definition of *book* changes

♦ ♦ ♦ *The Book*



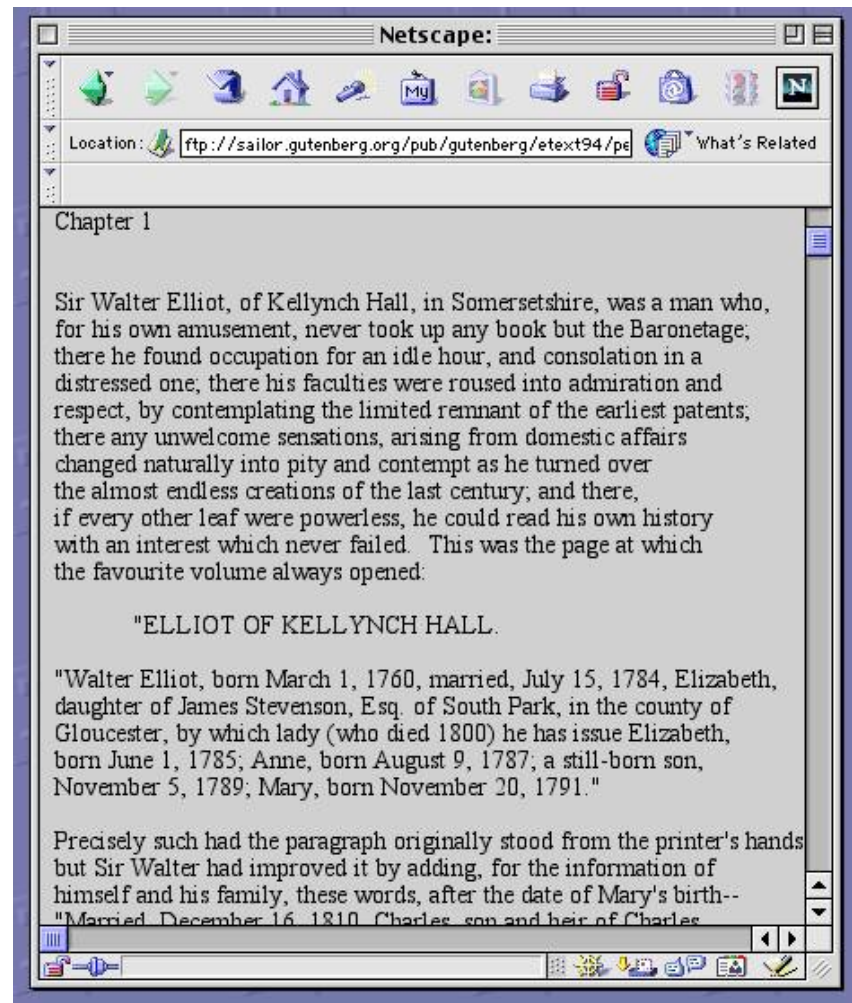
Books...



... and the books we actually read



Reading Jane Austen Online: Like touring Kent through a bombsight



Millenarian Visions

Predictions that Internet/Web will decentralize authority & permit ground-up social & political organization.

Cf the "virtual corporation," "smart mobs," moveon.org, etc., which are said to facilitate direct democracy



The Emergence of "The Collective Mind"

Cf. James F. Moore (Berkman Center), on "the global consciousness of the second superpower":

The Internet and other interactive media continue to penetrate more and more deeply all world society... The collective power of texting, blogging, instant messaging, and email across millions of actors cannot be overestimated. Like a mind constituted of millions of inter-networked neurons, the social movement is capable of astonishingly rapid and sometimes subtle community consciousness and action... the emergent democracy of the second superpower is alive with touching and being touched by each other, as the community works to create wisdom and to take action....

Two Spatial Models of Electronic Discourse -- Piazzas or Souks?



Verona's Piazza Bra



The New Jerusalem

Cyberspace: A world in which the global traffic of knowledge, secrets, measurements, indicators, entertainments, and alter-human agency takes on form: sights, sounds, presences never seen on the surface of the earth blossoming in a vast electronic night. . . . From vast databases that constitute the culture's deposited wealth, every document is available, every recording is playable, and every picture is viewable. The realm of pure information, filling like a lake...

— Michael Benedikt, *Introduction to Cyberspace*, MIT Press, 1991

Cyberspace: The Morning After

As access to the Web has widened, legitimate information has been subsumed by a deluge of vanity "home pages," corporate marketing gimmicks and trashy infomercials. Anyone who has ever used a search engine to find a specific piece of information on the Internet will cringe at the familiar litany of disappointment. It is impossible to know where information comes from, who has paid for it, whether it is reliable and whether you will ever be able to find it again. A student looking for information on the Internet about, say, World War II, cannot know whether a given "page" has been posted by a legitimate historian or by a Holocaust revisionist. Brian Hecht, "Net Loss," *The New Republic*, February 17, 1997

The Problem of "Pollution"

Racist & hate Speech

Mary Douglas: "Dirt is matter out of place."

Pornography/violence

Whereas pornography was once furtively glimpsed at dimly lighted newsstands or seedy adult theaters, today it is everywhere. It pours in over the Internet, sometimes uninvited, sometimes via eagerly forwarded links (Paris Hilton, anyone?)...*Time*, January 19, 2004

Dimensions of the Problem

2000 OCLC study: 7.1 million unique Web sites (projects to 11 million unique sites as of September, 2001). 68,000 "adult" sites (projects to 100,000 in 2002)

Lawrence and Giles, 1999: 1.5 billion pages in publically indexable Web; 1.5% of servers contain "pornographic" material (est. 100,000 servers now).

Search engines taken together reach 50 percent of indexable Web.

Issues of "Accessibility"

Much made of these figures. Cf Blaise Cronin expert report in CIPA case: "It is estimated that 100,000 sites contain child pornography."

Cf Ocala, Florida *Star Banner*, which quotes a spokesman for the US Customs Service.

[Pamela Paul asserts] that increased porn supply on the Internet (260 million pages of it, by her count) creates increased demand. Review of *Pornified*, NYT, 9/11/05

20 million view adult pages... are apparently hosted on sites in the United States or Canada. NRC study, 2002

Distribution of sites & traffic on the Net: "Winner Take All"

Distribution of User Volume Among Sites

<u>% of Sites</u>	<u>% of traffic</u>
	<u>All Sites</u>
.01	32.36
1	55.63
5	74.81
10	82.26
50	94.92

source: Adamic & Huberman, 2000

How Porn Sites are Used : An Exception to "Winner Take All"

Distribution of User Volume Among Sites

<u>% of Sites</u>	<u>% of traffic</u>	
	<u>All Sites</u>	<u>Porn Sites</u>
.01	32.36	1.4
1	55.63	15.83
5	74.81	41.75
10	82.26	59.29
50	94.92	90.76

source: Adamic & Huberman, 2000

Responses to Rotten Information

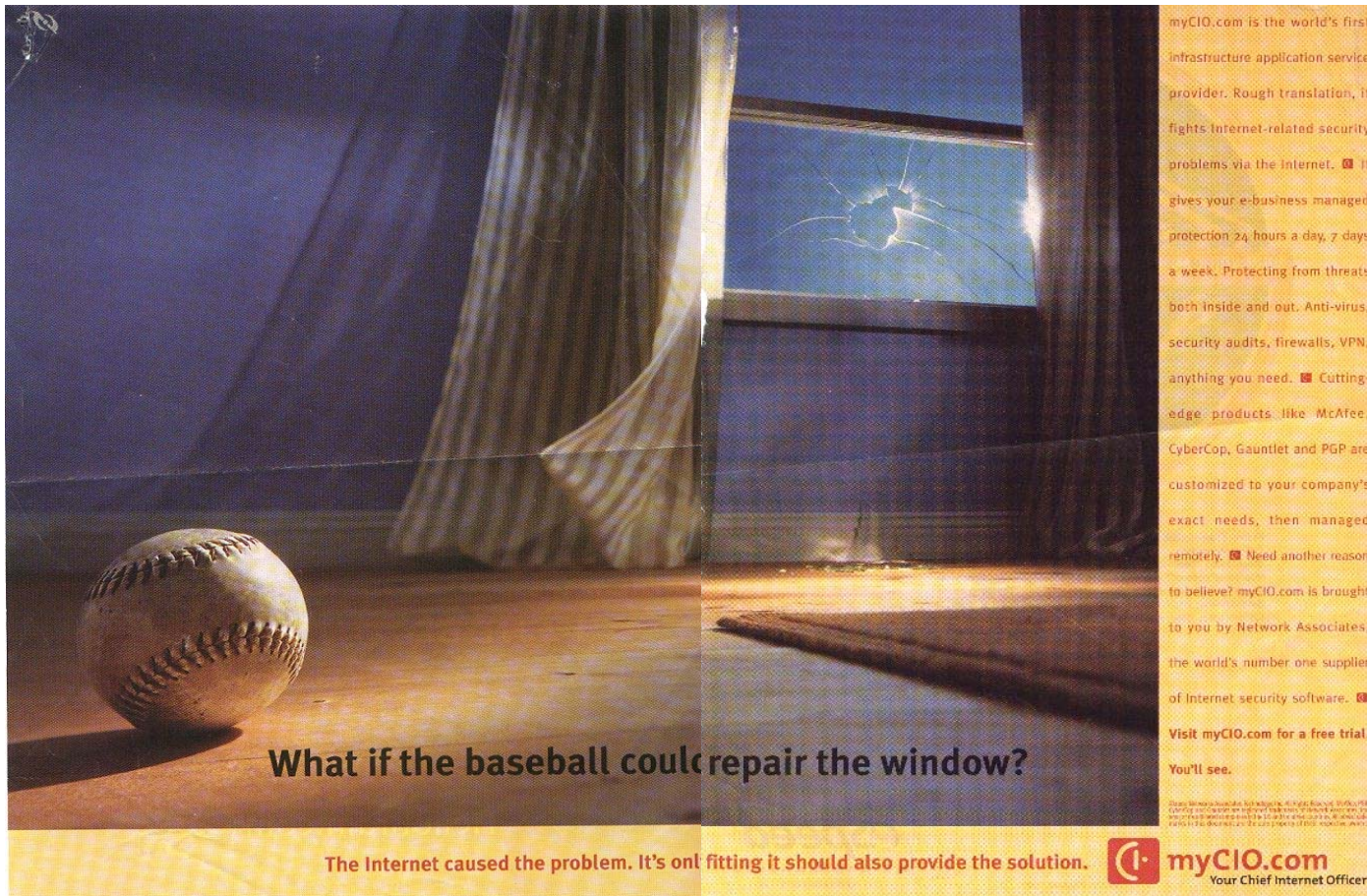
Institutional & organizational adaptations

Appeals to market forces

Legislative and legal remedies

Technological solutions

“The remedy for the abuse of digital technology is more digital technology.”



What if the baseball could repair the window?

The Internet caused the problem. It's only fitting it should also provide the solution.

myCIO.com is the world's first infrastructure application service provider. Rough translation, it fights Internet-related security problems via the Internet. ■ It gives your e-business managed protection 24 hours a day, 7 days a week. Protecting from threats both inside and out. Anti-virus, security audits, firewalls, VPN, anything you need. ■ Cutting-edge products like McAfee, CyberCop, Gauntlet and PGP are customized to your company's exact needs, then managed remotely. ■ Need another reason to believe? myCIO.com is brought to you by Network Associates, the world's number one supplier of Internet security software. ■ Visit myCIO.com for a free trial. You'll see.

© 2000 Network Associates, Inc. All rights reserved. All other trademarks are the property of their respective owners. Network Associates and the NA logo are registered trademarks of Network Associates, Inc. All other trademarks are the property of their respective owners.

myCIO.com
Your Chief Internet Officer

Feeding the Filtering Frenzy



Ad for McAfee's GuardDog software

Childrens' Internet Protection Act (CIPA)

Passed in 2001 -- 3rd attempt to regulate online porn

As condition for receiving E-rate discounts & grants under Lib. Services & Technology Act, mandates that schools & libraries must certify they are using a "technology protection measure" that prevents patrons from accessing "visual depictions" that are "obscene," "child pornography," or (for minors) "harmful to minors."

Bill provides for unblocking of sites for " bona fide research or other lawful purposes."

Challenge brought by ALA, ACLU, several other plaintiffs; heard by three-judge panel in March, 2002.

Problems with Filters: Overblocking

Filters block both:

"Relevant" sites:

Safe sex sites, sex studies sites, porn spoof sites, victims of porn" sites, gay magazines, site for "The Opposite of Sex," etc.

Also, sites or articles critical of filters

"Irrelevant" sites:

Latin music site, doll furniture site, obituary site, MTV site, low-rider site, autograph site, etc.

Edelman found overblocking of politicians' sites, cancer treatment facilities, rugby sites, site dealing with halitosis, etc.

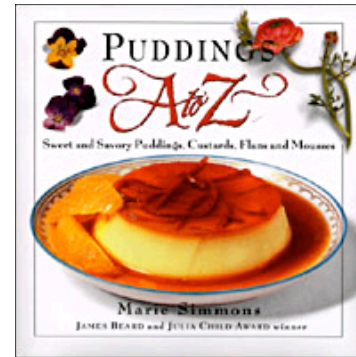
Filters in library use (government study): 6-10% of blocks are inappropriate.

Social & Institutional Responses to Internet Pornography

Libraries can regulate use by means of use policies, training and guidance, monitoring of logs, tap-on-the-shoulder policies, etc.

Rights of other patrons can be protected by privacy screens, etc.

Limits of Image Classification



Images found "Probably Objectionable" by WIPE (Wavelet Image Pornography Elimination / Detection).

Does the Web change everything?

"[Libraries] are simply declining to put on their computer screens the same content they have traditionally excluded from their bookshelves."

Solicitor General Ted Olson, arguing CIPA case before Supreme Court, 3/4/03

While Mr. [Paul] Smith emphasized the burden the law put on adults who have to request unblocking, Justice Stephen G. Breyer said: "What's the burden in asking? I grew up in a world where certain materials were kept in a special place."

Whereas pornography was once furtively glimpsed at dimly lighted newsstands or seedy adult theaters, today it is everywhere. It pours in over the Internet, sometimes uninvited, sometimes via eagerly forwarded links